EDITORIAL

ETDs and Digital Repositories— a Disciplinary Challenge to Open Access?

Charles B. Lowry

The University of Maryland Libraries have managed a repository using D-Space software for over two years, providing faculty a service for posting their research work and a foundation for moving the labor intensive management of paper dissertations and theses to the digital environment. Close cooperation with the Graduate School has been an essential feature of moving to a uniform requirement that theses and dissertations be presented in PDF format and posted in the Digital Repository at University of Maryland (DRUM). At an early stage, intellectual property issues began to emerge as an important policy dimension of managing DRUM—as they have for virtually any institution that gets into the digital repository business. I believe that these issues are indicative of some of the essential differences in the nature of information exchange among the disciplines and shape their reaction to open access. They arise particularly in the case of electronic theses and dissertations (ETDs) and have at least three dimensions:

> **I believe that these issues are indicative of some of the essential differences in the nature of information exchange among the disciplines and shape their reaction to open access.**

- For the university, the potential legal impact of providing access to copyrighted information included in ETDs by students
- For authors, the potential impact of posting on patent disclosure
- For authors, the potential impact of posting on later publication

I will not address issues of faculty-posted research but concentrate instead on ETDs that provide a sufficiently characteristic set of problems. Our experience with ETDs has surfaced deep fissures among disciplines in their attitudes about the electronic medium and has resulted in some unanticipated reactions to DRUM that reflect a certain

free-floating anxiety that caught us by surprise. It raises the fundamental need, indeed obligation, to educate our faculty and our graduate students about the realities they and we face. What must be achieved is a balancing act that honors two important academic traditions—the copyrights of authors and the research mission of broad access to the scholarly research output of the university. I believe that it is a balance we can achieve and still remain within the fundamental traditions of the academy.

Disciplinary Differences

In the UM campus debate, as elsewhere, the disciplinary point of view is distributed across a spectrum characterized roughly by the science and technology disciplines on the one end and the humanities and fine arts on the other, with the social and behavioral sciences falling somewhere in between. The two extremes are hardly surprising but still illustrative. Scientists and engineers readily accept, even insist on, the pre-publication posting of dissertations. Anything less is viewed as a breach of the scholarly canon of sharing new knowledge quickly and building on it. When they have concerns, they resolve them directly and pragmatically. For instance, at the University of Maryland if there are patent disclosure issues, they ask only that a brief embargo of no more than a year be provided. In the case of publishers who decline articles that are drawn from dissertations, they ask for a year embargo and insist that it be based on explicit publisher policy such as that of the American Chemical Society.

At the other end of the spectrum there is a high level of concern about posting, in some cases any posting at all. The arguments are varied and often anecdotal—that some unidentified book publishers decline posted dissertations, that posted dissertations are fair game and will be ripped off by unscrupulous scholars, and that posting itself places too much burden on students to clear copyright of materials they use in their dissertations. I have heard the case made for embargoes of 20 years and more. I am not making an invidious comparison among the disciplines but merely paraphrasing the positions taken in our debate, neither am I unsympathetic to the concerns of the humanities and fine arts since I am a humanist by training. Moreover, the UM Libraries have partnered with the College of Arts and Humanities in founding the Maryland Institute for Technology in the Humanities (MITH) with an NEH challenge grant, which is committed to the migration of humanities scholarship to the digital world.

I believe these rather stark differences among disciplines are the result of the relative penetration of different disciplines into the electronic medium and the length of their experience with it. This point has been made quite recently and emphatically in "Our Cultural Commonwealth: the Report of the American Council of Learned Societies' Commission on Cyberinfrastructure for Humanities and Social Sciences," www.acls.org/cyberinfrastructure / cyber_report.htm. I believe the differences will dissipate with time—but they are very real now and not to be ignored. To date at Maryland, we have dealt with these concerns on an ad hoc basis, restricting access to less than a

> **I believe these rather stark differences among disciplines are the result of the relative penetration of different disciplines into the electronic medium and the length of their experience with it.**

dozen dissertations, a small number when one considers that currently there are over 3,000 documents in DRUM, and two-thirds are ETDs.

Copyrights and Their Uses

Within the current copyright regime, the higher education tradition in the United States has been to treat students as we do faculty with respect to copyright. Thus, student authors—unequivocally—own copyrights to their works; but at the same time, by policy, the university may require freedom to use the works. In addition, as a condition of graduation, universities have the right (and obligation) to archive theses and dissertations physically or electronically for purposes of research and education. Of importance, too, is the typical university policy requirement that students obtain permission to use copyrighted materials they incorporate into their work. The corollary is that the university is not obligated to protect them from violations. Such practices are well instantiated in policy and tradition. As an example, University of Maryland students are required by the Graduate School to agree to the following terms when submitting their works:

> I hereby grant the University of Maryland, College Park, a non-exclusive, fully paid, irrevocable right and license to host, cache, route, transmit, archive, reproduce, distribute, perform, display, and reformat My Manuscript [dissertation and abstract or thesis and abstract], in whole or in part, in and from an electronic format, as of the submission date stipulated in the submission of this document and approved by the University of Maryland. I understand that I maintain the right to request that the University of Maryland restrict electronic access to my document for a specified time, through procedures outlined in the *University of Maryland Graduate Catalog and Handbook for Graduate Students and Graduate Advisors*. I warrant that I have submitted to the University a completed Thesis and Dissertation Electronic Publication Form. I understand that the University of Maryland will submit My Manuscript to PQIL® [ProQuest Information and Learning Company] Dissertations Publishing for inclusion in the publications *Dissertation Abstracts International* or *Masters Abstracts International*. I hereby grant PQIL a non-exclusive right to host, cache, route, transmit, archive, reproduce, distribute, perform, display, and reformat My Manuscript, in whole or in part, in any format whether in existence now or developed in the future, including but not limited to microfilm and any electronic formats. Other publication rights may be granted as I choose. I represent and warrant that (a) My Manuscript is my original work; (b) any third party content included in My Manuscript is either in the public domain, constitutes a fair use under copyright law, or is included with the written permission from the owner of copyright in such third party content; (c) excluding third party content, I am the sole owner of My Manuscript and have the authority to grant these licenses, and (d) does not violate the copyright, privacy or publicity, or any other personal, proprietary or intellectual property rights of any third party and is free of scandalous, libelous, and/or unlawful matter. I agree to supply the University of Maryland with copies of the permissions I have obtained from third parties to use their copyrighted content and acknowledge that ProQuest and the University of Maryland have the right not to distribute My Manuscript if, in their reasonable judgment, either believes that My Manuscript violates any of these representations and warranties. I acknowledge that I retain copyright and ownership rights in My Manuscript and have the right to exercise any of my rights under copyright in My Manuscript, including for example, publishing a book and/or articles based on My Manuscript.[1]

For a century and more, the archival access to these works has undergirded further research and publication by other students and scholars. If we restrict the access to ETDs, they become gray literature, inaccessible and fundamentally invisible to the scholarly record. Historically, access has been provided in print through a robust system of ILL nationally and internationally and by vigorous collecting of dissertations, in particular, by research libraries and the Center for Research Libraries. In addition, until the late 1990s the Library of Congress collected and archived U.S. dissertations, a responsibility it has ceded to Proquest (the former UMI). Thus dissertations, if not theses, have been broadly available and located through the access provided by *Dissertation Abstracts* and *Masters Abstracts* and their online derivatives, as well as through library catalogs. Proquest, for its part, has responded to pressures to limit access at the author's request—for example, embargoes and permanent restriction of access to everything but the bibliographic description and abstracts. In addition, Proquest redacts information from ETDs where there is doubt that the author obtained copyright clearance. On the other hand, the appearance of digital dissertations and theses does not expand awareness of these works, but it does make them more easily accessible—a fundamental goal of libraries and universities in establishing digital repositories.

> **If we restrict the access to ETDs, they become gray literature, inaccessible and fundamentally invisible to the scholarly record.**

At the same time, graduates turn theses and dissertations into books, journal articles, patents, and everything from performance to software. In the hardcopy world of print, publishers never took the position that the appearance of a dissertation through any of these means was prior publication. It is worth reminding ourselves that under the 1976 copyright law, once in fixed form (print or electronic), dissertations are copyrighted works—that is, they are ipso facto published. With the appearance of various forms of online posting and repositories, some publishers have taken the position that posting a digital dissertation *is* prior publication—but only a small fraction—and they do so for reasons that have nothing to do with distribution of scholarship. Humanities and social science students are principally those who turn dissertations into books—usually after extensive revision. There is evidence that electronic release of early versions (such as dissertations) leads to greater sales of such books. Thus, having an electronic work made available on the Internet and showing a publisher a large amount of electronic use of that work may help land a book contract. In the case of journal articles, in some disciplines (for example, high energy physics and computer science), pre-prints (really pre-publication) are a firmly fixed practice that has migrated seamlessly to the world of the Internet. This principally occurs in the science and technology disciplines; and the later appearance as a double-blind reviewed journal article is considered the final stage of validation of the research, giving it the disciplinary stamp of approval. Journal publishers, by and large, have come to realize this, and the trend is to allow pre-posting and often

> **We in the academy who produce the knowledge that publishers exploit for profit should encourage this trend of allowing pre-posting by all means possible.**

additional posting after publication. A few, though, have explicitly stated they will not publish previously posted work. Moreover, the position of even restrictive publishers such as these is in flux and moving toward allowing pre-posting. We in the academy who produce the knowledge that publishers exploit for profit should encourage this trend of allowing pre-posting by all means possible.

Resolving the Dilemmas

In many, perhaps the majority, of cases, the dissertation or thesis is and will be the only version of the work that ever appears. In other cases, the appearance will have no effect on the author's later ability to publish a version(s) of the work. On the other hand, in some cases digital posting may compromise the future prospects for the author to publish. We should not jeopardize access in the former cases to protect in the latter. We can look to the community for practices that are emerging. For instance, there are 38 (by my latest count) members of the Association of Research Libraries that have established digital repositories of dissertations. Half of these are also members of AAU. Twenty of them permit embargoes that range from six months to five years. But these embargoes are granted principally by virtue of special request and for cause. All this repository activity has the added advantage of opening access to search engines, expanding access on the open Internet. In the end, the dilemmas we face will be resolved not in the technology framework but in the political and policy frameworks. At Maryland, in concert with the Graduate Council and dean of the Graduate School, we have developed a solution that is not atypical, in spite of any institutional idiosyncrasies, and that has the following four elements:

Access Issues with an Embargo

There are two ways to manage limited campus access to embargoed ETDs that we have investigated at UM. Creating campus directory authentication allows the user to log in and can be limited to the remainder of the current session. Alternatively, using an IP address would restrict use to campus. Functionally, the difference between the two is that authentication allows a member of the campus community to have access to the digital object from anywhere, whereas IP address allows access to the digital object from a computer with a UM IP address, regardless of who the person may be. Thus the former method is truly campus community access only, but the latter allows visitors to the campus to have access to the digital objects and prevents access by our own students and faculty from off-campus. Our current policy does not allow this kind of access to ETDs when an embargo has been requested by the author, but we will bring a proposal to do this before the Graduate Council in the near future. Whether limited access to PDFs is allowed, we make the bibliographic information and abstracts of embargoed ETDs available on DRUM.

Term Limits for an Embargo

Our current practice is to suppress access to the digital objects based on the issue date of each ETD and to remove the suppression based on the issue date. Each ETD automatically becomes available when the time restriction expires. Providing open-ended

user-selected options regarding the length of the time restriction for any given ETD is not desirable from the administrative point of view. We have reached agreement to implement routinely three embargo periods—one year, six years, and indefinite—but the default is for immediate access. If the concern is that people need to publish a book for tenure and promotion, then the longer period allows them to accomplish that by the time the tenure clock expires. However, in special cases the author may request an indefinite embargo under the policy.

> **Providing open-ended user-selected options regarding the length of the time restriction for any given ETD is not desirable from the administrative point of view.**

> You may place an indefinite embargo on your work. In this case, a description of your work including your name, the title, your advisor's name, and the abstract will be available via ProQuest's *Digital Dissertations* and DRUM, but the actual electronic file will be embargoed indefinitely. This option requires the written approval of the Dean of the Graduate School. This restriction can be lifted at the request of the author at a later date.[2]

It is our intention to regularly ask authors who have been granted an indefinite embargo to remove it. We expect this provision to be exercised rarely.

Review and Approval Protocols

In the case of a one or five year embargo, the justification for the embargo must be explicit and real so that the Graduate School approval is based on facts not fears. The Graduate School has established a systematic review and application system for authors who have a demonstrated need for delay. For this to be effective and judicious, we recognize that a formal review process must be established similar to that at other institutions that grant embargoes. Above all, advisors are involved in the request process, and there is, as we have seen, a proactive decision by the dean of the Graduate School. For their part, faculty advisers are required to counsel with advisees on the options concerning access and to sign a request for any of the three embargo types. Finally, we think it is helpful to provide students and faculty guidance with respect to journal and book publishers' policies concerning pre-posting of digital dissertations. Currently available Web sites such as SHERPA/ROMEO and AcqWeb do this.[3]

Hardcopy Access

We accept nothing less than full access to the print dissertations as soon as they arrive and are processed for the Special Collections. In reality, this does not occur until six to eight months after the dissertation is submitted simply because the hardcopy comes from Proquest. We no longer acquire circulating or ILL copies and fulfill our obligation for providing this archival access to the larger academic community from Proquest and DRUM. Thus, the embargo period might have the impact of limiting non-UM use of our dissertations unless the researcher is willing to come to campus to use it in the library or to pay for hardcopy ILL on a cost-recovery basis by printing on demand. We do provide for the latter case in the Graduate School policy:

These choices [embargoes] only affect the *electronic* distribution of your thesis or dissertation document. A non-circulating copy of your thesis or dissertation will be available for consultation in Hornbake Library's Maryland Room, and print copies of your document will be made available upon request to researchers through Inter-library Loan.[4]

Conclusion

In the end, the real reason for tackling the disciplinary differences that repository posting or archiving entail is that it is the right thing to do. Solving the attendant problems provides the opportunity for increasing disciplinary awareness of the opportunities inherent in the electronic environment, allows us to expand access, and ensures the continuation of basic institutional values of libraries and higher education. We fully expect to learn from this experience and also use it as an opportunity to educate our community. We know that our practice concerning ETDs presents a moving target and that in time the practice will inevitably change. Indeed, it would be a great irony if we were to take actions that now made digital dissertations less available than print-only dissertations historically have been.

> In the end, the real reason for tackling the disciplinary differences that repository posting or archiving entail is that it is the right thing to do.

Charles B. Lowry is dean of libraries, University of Maryland College Park, MD; he may be contacted via e-mail at: clowry@umd.edu.

Notes

1. "Proquest / University of Maryland Agreement for Submission of Theses and Dissertations" (draft document).
2. "Electronic Publication Form" (draft document).
3. There are two sites, at least, that provide authors information about publisher policies concerning posting. For journal publishers, the University of Nottingham supports the SHERPA/RoMEO Web page "Publisher Copyright Policies and Self-archiving," http://www.sherpa.ac.uk/romeo.php?all=yes (accessed July 24, 2006). Vanderbilt University's AcqWeb provides a handy "Directory of Publishers and Vendors" that provides information for contacting specific publishers with inquiries, http://acqweb.library.vanderbilt.edu/pubr.html (accessed July 24, 2006).
4. "Publishing Your Thesis or Dissertation: ProQuest Digital Dissertations©& DRUM, the Digital Repository at the University of Maryland" (draft document).